构建安全保障体系，护航人工智能新基建发展

魏薇　牛金行　景慧昀

（中国信息通信研究院安全研究所，北京 100191）

摘要：人工智能作为新基建的核心组成部分，在垂直行业融合应用日趋广泛和持续深化，正在催生出众多智能化新产品和新业态，为经济社会高质量发展注入新动能。亟需构建安全保障体系，全面应对人工智能新基建风险挑战，推动国家战略顺利实施和经济社会转型发展。立足于推进我国人工智能新基建健康发展，基于文献研究方法，结合产业问题调研，归纳总结人工智能新基建发展过程中面临的国际博弈、技术局限、安全问题、治理体系等诸多挑战，研提安全保障体系建设的思路建议，旨在为智能产业稳步发展提供参考借鉴。

关键词：人工智能；新基建；安全治理；高质量发展

中图分类号：TP391.7　　**文献标识码**：A

引用格式：魏薇，牛金行，景慧昀．构建安全保障体系，护航人工智能新基建发展[J]．信息通信技术与政策，2021，47(5)：11-14.

doi：10.12267/j.issn.2096-5931.2021.05.003

0　引言

当前，受新冠肺炎疫情防控、国际竞争博弈等影响，我国传统拉动经济增长的外贸、投资和消费“三驾马车”均受到不同程度的影响，经济面临下行压力[1]。国家多次提出加快新型基础设施建设，国家部委和地方政府积极出台“新基建”相关指导意见和投资计划。相较于传统基建，“新基建”发展空间巨大，带动效应明显，正在成为我国稳投资、调结构、扩内需的经济发展新引擎。人工智能作为“新基建”的核心板块之一，其发展备受瞩目。在新冠肺炎疫情防控期间，人工智能企业纷纷开放算力、算法和数据等基础设施资源，助力提升传统行业智能化水平，为对冲疫情影响、促进产能恢复发挥了重要支撑作用。

人工智能作为“新基建”，更加强调其基础设施属性，主要包含以下特征：一是基础支撑性，人工智能新基建着力推进人工智能计算、存储、网络等硬件基础设施和框架、算法、数据等软性基础设施建设，提升数字基础设施支撑能力；二是公共服务性，人工智能基础算力、行业算法、数据资源和通用软件将更多以平台服务或者授权许可方式来提供，便利各行业、各领域获取普惠性的智能化服务；三是溢出带动性，人工智能新基建会加快人工智能技术的融合应用推广，推动传统基础设施智能化升级改造，赋能传统行业高质量转型发展，成为培育壮大数字经济的重要支撑。

1　人工智能新基建推进面临的挑战

数字经济时代，世界主要国家大力促进新兴技术发展，抢抓新一轮科技革命和产业变革的重大机遇，以5G、人工智能为代表的新型基础设施建设成为推进信息技术产业化的关键支撑，以及提升国家核心竞争力的重要抓手。面对国际竞争新形势、技术推广新困境、泛在安全新风险以及安全治理新问题，人工智能新基建的推进实施存在很多良机，同时也面临诸多挑战。

1.1　从国际形势看，人工智能新基建发展主导权受到威胁

当前，我国人口红利逐渐消退，传统制造业有加速向外转移趋势，而具备高附加值的科技和服务产业又

由发达国家把持[2]。为避免落入“中等收入陷阱”,我国亟需把握以人工智能为代表的新一轮科技革命战略机遇,寻求新的经济增长点。目前,我国在计算机视觉、智能语音等人工智能应用层面处于全球领先水平,但在算力芯片、学习框架、基础算法等人工智能基础设施方面,对外依赖较为严重。近年来,以美国为首的发达国家为锁定其技术优势和主导地位,持续加大基础环节管控,遏制我国人工智能行业发展。自 2019 年至今,美国政府已将我国数十家知名人工智能企业列入“实体清单”,限制人工智能相关基础软硬件出口。后疫情时代,部分国家逆全球化思维呈现加速发展态势。美国加入全球人工智能合作伙伴组织(Global Partnership on Artificial Intelligence,GPAI)并推动发布联合声明,强调以符合“人权、自由和共同的民主价值观”的方式支持 AI 开发和使用,这将在一定程度上限制我国人工智能技术在全球范围内的发展和应用。错综复杂的国际形势给我国人工智能新基建推进带来新的风险挑战,影响自主发展能力,制约产业全球推广。

1.2 从技术特征看,人工智能新基建推进本身存在复杂性

人工智能技术快速发展演进,通用智能的实现方法尚不清晰,以海量数据和深度学习为驱动的技术途径虽然取得显著成果,但仍存在数据强依赖性、算法弱鲁棒性、技术高复杂性等诸多瓶颈,制约人工智能新基建推进应用。一是数据强依赖性制约垂直行业应用推广。人工智能充分赋能行业需要依赖行业内海量优质的应用场景数据,但在传统的工业制造、农业生产、卫生医疗等具体行业内,往往存在数字化转型程度不高以及数据开放共享不足等问题,对数据资源的高要求将限制人工智能有效推广[3]。二是算法弱鲁棒性较难适应复杂场景应用。现阶段深度学习算法存在鲁棒性弱的问题,在具体应用场景的开放动态环境中,算法决策可能会产生意料不到的错误,导致人工智能不太适用于工业控制、能源输配等安全可靠性要求高的场合。三是技术高复杂性限制新基建推进进程。人工智能具有多学科交叉、高度复杂性特征,导致专业人才无法及时供给,缺口较大,而人工智能新基建推进需要大量既懂人工智能技术、又有行业领域知识的复合型人才,此类人才目前更为稀缺。根据教育部门测算,我国人工智能人才缺口超过 500 万,供需比例严重失衡,这将限制人工智能新基建的快速推进。

1.3 从安全风险看,人工智能新基建面临多层次安全挑战

新基建战略推进将促使人工智能成为具有基础支撑性、公共服务性等特征的社会公共品,进而对其安全性和可靠性水平提出更高要求。然而,当前阶段人工智能技术仍不断演进完善,且安全防御理论和技术处于探索初期,尚无法有效应对愈加复杂、多维度、多层次安全挑战。一是人工智能数据和算法安全风险更加突出。新基建战略的实施将加速人工智能技术以开源开放的平台、算法包、模型库等形式,向社会提供开放共享的普惠性服务。但是,受限于人工智能数据安全保护技术和机制尚不健全成熟,用户上传至平台的数据面临被第三方窃取、隐私泄露等风险。此外,第三方预训练人工智能算法包面临恶意提供者嵌入新型后门和木马的安全风险[4]。此类安全风险非常隐蔽,用户难以检测发现,给后续人工智能应用带来安全隐患。二是人工智能外部安全攻击威胁更加严峻。新基建推动人工智能基础平台向行业应用平台加速演进,人工智能技术在交通、医疗、金融等行业应用持续深化,广泛赋能云侧平台和端侧设备。由此,人工智能产品外部攻击面不断延展,受攻击的可能大大增加,不良分子可从智能终端等安全防护薄弱环节实施攻击,通过劫持终端入侵系统,进而威胁云侧人工智能基础设施安全。

1.4 从治理体系看,人工智能新基建带来安全治理新问题

从国家发展和改革委员会对于“新基建”概念界定来看,基于人工智能的融合基础设施成为未来经济社会发展的重要支撑,将催生出大量的经济新业态和商业新模式[5]。但是,融合基础设施以及构建在其上的智能应用的安全监管工作会涉及到政府多个管理部门,原有基于传统行业界限的部门管理职责划分将无法适用新技术发展。例如,智能网联汽车的安全监管会涉及国家多个部委,各部门会结合自身原有职责从不同维度和切入点开展监管工作,难免出现交叉监管、监管空白等机制体制问题。

另外,人工智能深度学习算法的不透明性、难以追责等问题,给现有安全监管技术带来极大挑战,如何对人工智能安全风险实时监测和及时处置,成为安全治

理机制建设的重点内容。随着人工智能技术的快速发展和应用场景的不断泛在,亟需加强部门统筹协同、创新监管机制手段,以保障人工智能新基建的顺利推进。

2 构建人工智能新基建安全保障体系思路建议

新基建为国家人工智能战略落地注入新动能,成为我国人工智能技术产业补短板、强弱项、促应用的新机遇。我国人工智能新基建面临着发展和安全、国内和国际、技术和治理等多维度挑战,为保障我国人工智能新基建顺利实施,亟需统筹考虑,加快推动人工智能核心技术创新、治理体系构建、安全能力提升以及国际合作深化,健全完善我国人工智能安全保障体系。

2.1 提升人工智能核心技术创新能力

一是立足自主创新补足技术短板。针对先进制程芯片、基础设计软件等“卡脖子”的关键核心技术环节,在新基建战略推进过程中,注重发挥国家和地方政府专项产业基金引导作用,鼓励龙头企业聚集产业链上下游公司联合开展短板技术攻关和国产化技术应用,缩短核心技术研发攻关周期。二是加强海外人才引进和本土人才培养。针对人工智能新基建面临的高端人才短缺困境,加速引入海外知名人才中介服务机构,创新海外高端人才柔性引进机制,授予国有企业更大的海外高端人才薪酬激励自由度。创新校企联合教育模式,以新基建需求为导向,增强交叉领域复合型人才培养。三是利用全球创新资源提升技术创新能力。利用我国超大规模市场虹吸效应以及持续优化的外资营商环境,吸引面向我国市场的海外人工智能软硬件企业加速在我国设立设计研发基地,增强我国人工智能产业整体技术创新能力。

2.2 建立多方参与的人工智能安全治理体系

一是政府发挥安全治理主导作用。整合多学科力量,加强人工智能新基建面临的法律、伦理、社会等方面的突出问题研究,建立涵盖政府、企业和社会各方的人工智能安全责任体系[6],制定出台人工智能安全治理政策指南,构建人工智能应用分级分类安全治理机制。二是企业积极践行安全治理主体责任。企业作为人工智能技术研发与应用的主要力量和一线实践者需承担治理主体责任。企业建立内部人工智能安全治理规范和制度,设立人工智能治理机构,从法律、伦理、社会等视角加强对各项人工智能应用的合理性审查,探索并实施算法偏见、算法黑箱、算法弱鲁棒性、数据隐私等问题的技术解决措施,加强出厂前人工智能产品安全性核实验证。三是社会加强安全治理监督。加快培育人工智能安全检测咨询服务机构,构建人工智能安全检测验证公共服务平台,依托行业联盟建立人工智能安全投诉举报、核实验证、公开曝光渠道。

2.3 健全人工智能安全技术保障能力

一是完善人工智能安全技术标准体系。制定出台人工智能安全标准体系框架,构建人工智能安全标准推进路线图,加快研制人工智能数据安全、算法模型安全、产品和应用安全、安全检测评估等亟需标准。在人工智能新基建先导应用领域,加强人工智能安全标准宣贯和试点验证。二是加强人工智能安全技术和产品研发。新基建规划加强了人工智能安全项目布局,针对人工智能新基建面临的突出算法和数据安全风险,鼓励人工智能企业和网络安全企业充分发挥各自优势,通过联合研发等方式,开展人工智能安全防御技术和产品攻关。三是加快人工智能安全保障技术平台建设。通过国家重点研发计划等专项资金,加快人工智能安全检测、监测预警、应急处置、追踪溯源等技术平台建设。

2.4 促进人工智能国际互信合作发展

一是增进人工智能发展国际互信。依托国际政府合作组织、标准化协会等,广泛分享我国人工智能新基建推进、融合应用以及安全治理工作的有益尝试和成功经验,积极参与或主导人工智能国际标准制定工作,推动形成人工智能发展和治理国际共识,增进国家间包容互信。二是加强人工智能新基建国际合作。鼓励国内人工智能龙头企业持续加强全球布局,积极开展国际业务,以东盟、亚太地区以及“一带一路”沿线国家为合作重点,努力向欧美市场拓展延伸,推进人工智能新基建全球化发展,为我国人工智能发展构建广泛生态系统和宽松国际环境。

3 结束语

在国家加快推动“新基建”背景之下,人工智能将进入高速发展阶段,其溢出带动效应更为凸显,为传统行业数字化转型和经济社会智能化发展提供强劲引擎。立足当下,着眼未来,需高度重视人工智能新基建

发展过程中的安全风险和问题挑战,开展应对举措前瞻研究,探索构建安全保障体系,为人工智能新基建的健康发展保驾护航。

参考文献

[1] 中国工商银行现代金融研究院课题组, 周月秋, 樊志刚. 后疫情时代全球经济展望[J]. 现代金融导刊, 2021(2):42-45.
[2] 周亚敏. 全球价值链中的绿色治理: 南北国家的地位调整与关系重塑[J]. 外交评论, 2019(1): 49-80.
[3] 中国电子信息产业发展研究院. 中国"新基建"发展研究报告[R]. 北京, 2020.
[4] 张思思, 左信, 刘建伟. 深度学习中的对抗样本问题[J]. 计算机学报, 2019(8): 15.
[5] 石梦. 新基建对经济发展的作用探讨[J]. 中国集体经济, 2021(3): 19-20.
[6] 郭亚军, 卢星宇, 张瀚文. 人工智能赋能信息无障碍: 模式、问题与展望[J]. 情报理论与实践, 2020(8): 57-62.

作者简介:

魏薇 中国信息通信研究院安全研究所正高级工程师,主要从事信息安全、数据安全、人工智能安全、安全评估等方面的研究工作

牛金行 中国信息通信研究院安全研究所高级工程师,主要从事人工智能安全、信息安全等方面的研究工作

景慧昀 中国信息通信研究院安全研究所高级工程师,主要从事人工智能安全、信息安全等方面的研究工作

Build security system to escort the development of new infrastructure for artificial intelligence

WEI Wei, NIU Jinhang, JING Huiyun

(Security Research Institute, China Academy of Information and Communications Technology, Beijing 100191)

Abstract: As a core component of the new infrastructure, artificial intelligence (AI) is more and more widely applied in vertical industries and continues to deepen, giving rise to many intelligent new products and new business models, and propelling for a high-quality economic and social development. It is urgent to build a security safeguard system to comprehensively address the challenges of AI as the new infrastructure and promote a smooth implementation of national strategies and socio-econmic transformation and development. To promote the healthy development of AI new infrastructure in China, this paper uses literature research methods, combining with industrial issues research, summarizing the challenges such as international game, technical limitations, security issues, governance system, and puts forward suggestions for the construction of the safety guarantee system, aiming to provide a reference for the development of intelligent industry.

Keywords: artificial intelligence; new infrastructure; security governance; high quality development

(收稿日期:2021-04-15)